

Wasserstein metric between a discrete probability measure and a continuous one

Weihua Yang and Xia Wang

December 16, 2021

Abstract

The paper considers Wasserstein metric between the empirical probability measure of n discrete random variables and a continuous uniform one on the d -dimensional ball and give the asymptotic estimation of their expectation as $n \rightarrow \infty$. Further We considers the above problem on a mixed process, i.e., n discrete random variables are produced by the Poisson process.

Keywords: Wasserstein metric; optimal matching; from discrete to continuous; random variable; Poisson process.

1 Introduction

Article [1] studied Ollivier curvature of random geometric graphs, a key step in which is building up the estimation of Wasserstein metric between the empirical probability measure of n discrete random variables and a continuous uniform one on the d -dimensional ball and authors applied results in [5], but which are building on $[0, 1]^d$ versus Ollivier curvature is built on balls, so the proof process in [1] seems tedious and vulnerable. Another hand, Lattice method in statistical mechanical approach [6] often involve similar notation and convergence from a discrete physical quantity to a continuous one, which should exist close links with the convergence from a discrete probability to a continuous one. All inspire us to study approach problem between the discrete probability and the continuous one, in this article we choose the Wasserstein metric as the approach measure, by which we hope to build a bridge in studying the relation between a discrete quantity to a continuous one in more scenes.

2 Preliminary Estimation

Definiton 1. Let $X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_n$ be independent uniformly distributed random variables on the d -dimensional $B(0; 1) = \{x \in R^d, \|x\| \leq 1\}$, $d \geq 2$, where $\|\cdot\|$ is a norm on R^d . The random variable $M_n^d := \inf_{\sigma} \sum_{i=1}^n \|X_i - Y_{\sigma(i)}\|$ denotes the optimal matching between X_1, X_2, \dots, X_n and Y_1, Y_2, \dots, Y_n , where σ takes over all permutations of $\{1, 2, \dots, n\}$.

By dual principle [2, 3], one has

$$M_n^d = \sup_{f \in \mathcal{L}_1} \sum_{i=1}^n (f(X_i) - f(Y_i)) = \sup_{f \in \mathcal{L}_1} \left| \sum_{i=1}^n (f(X_i) - f(Y_i)) \right|,$$

where the set of Lipschitz functions $\mathcal{L}_1 = \{f : B(0; 1) \rightarrow R; |f(x) - f(y)| \leq \|x - y\|, \forall x, y \in B(0; 1), f(0) = 0\}$. Notice every Lipschitz function in \mathcal{L}_1 may extend to one in

$$\mathcal{L} = \{f : R^d \rightarrow R; |f(x) - f(y)| \leq \|x - y\|, \forall x, y \in R^d, f(0) = 0, \|f\|_{L^\infty} \leq 1\}, \quad (1)$$

so $\mathcal{L}_1 = \mathcal{L}|_{B(0;1)}$. The following Lemma 1 gives a upper and lower bound estimation for the expectation $E(M_n^d)$.

Lemma 1 (optimal-matching). *For the above optimal matching, one has for the dimension $d \geq 2$*

$$1 - \frac{1}{d+1} \leq \liminf_{n \rightarrow \infty} \frac{E(M_n^d)}{n^{1-\frac{1}{d}}} \leq \limsup_{n \rightarrow \infty} \frac{E(M_n^d)}{n^{1-\frac{1}{d}}} \leq 2d + 5 + 8\sqrt{2}.$$

Proof. we put the proof in appendix. The method is essentially from [5], we make some improvement and modification for extending it to random variables on balls.

3 Main results and proofs

The following are results on the Wasserstein metric between empirical and uniform measures on the d -dimensional ball. In general, the Wasserstein metric between two probability μ_1, μ_2 is given by the following definition.

Definiton 2. *Let μ_1 and μ_2 be Borel probability measures on a compact metric space (\mathcal{X}, d) and let $\Gamma(\mu_1, \mu_2)$ denote the set of joint probability μ between μ_1 and μ_2 , whose marginals are μ_1 and μ_2 respectively. The Wasserstein metric is defined by*

$$W(\mu_1, \mu_2) = \inf_{\mu \in \Gamma(\mu_1, \mu_2)} \int_{\mathcal{X} \times \mathcal{X}} d(x, y) d\mu(x, y).$$

By duality principle (Kantorovich Dual Theorem) [7], Wasserstein metric can be expressed as

$$W(\mu_1, \mu_2) = \sup_{f \in \mathcal{L}_1(\mathcal{X})} \left(\int_{\mathcal{X}} f(x) d\mu_1(x) - \int_{\mathcal{X}} f(y) d\mu_2(y) \right),$$

where $\mathcal{L}_1(\mathcal{X})$ denotes the set of Lipschitz functions about metric on \mathcal{X} with the coefficient 1. From the above duality formula, we may further assume $\mathcal{L}_1(\mathcal{X})$ with the additional condition that is any $f \in \mathcal{L}_1(\mathcal{X})$ satisfying $f(0) = 0$.

Notice: In below all metrics are induced by some norms. To keep uniform math notation with most original articles, we still use ' d ' expressed the metric in a space that shouldn't arise confusion with the dimension notation.

Theorem 1. *Let X_1, X_2, \dots, X_n be independent uniformly distributed random variables on the d -dimensional ball $B(0; 1)$, let m_n^d denote the empirical measure*

$$m_n^d(y) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{y=X_i}$$

and μ^d the uniform measure on $B(0; 1)$. Then as $n \rightarrow \infty$,

$$E[W^d(m_n^d, \mu^d)] = O(n^{-\frac{1}{d}}), d \geq 2.$$

Proof.

$$\begin{aligned}
 W^d(m_n^d, \mu^d) &= \inf_{\mu \in \Gamma(m_n^d, \mu^d)} \int_{B(0;1) \times B(0;1)} d(x, y) d\mu(x, y) \\
 &= \sup_{f \in \mathcal{L}_1(B(0;1))} \left(\int_{B(0;1)} f(x) dm_n^d(x) - \int_{B(0;1)} f(y) d\mu^d(y) \right). \tag{2}
 \end{aligned}$$

Let Y_1, Y_2, \dots, Y_n be independent uniformly distributed random variables on $B(0; 1)$, then

$$\int_{B(0;1)} f(y) d\mu^d(y) = E[f(Y_i)], i = 1, \dots, n.$$

So

$$\begin{aligned}
 W^d(m_n^d, \mu^d) &= \sup_{f \in \mathcal{L}_1(B(0;1))} \left(\int_{B(0;1)} f(x) dm_n^d(x) - \int_{B(0;1)} f(y) d\mu^d(y) \right) \\
 &= \frac{1}{n} \sup_{f \in \mathcal{L}_1(B(0;1))} \left(\sum_{i=1}^n (f(X_i) - E[f(Y_i)]) \right) \\
 &= \frac{1}{n} \sup_{f \in \mathcal{L}_1(B(0;1))} \left(\sum_{i=1}^n E[f(X_i) - f(Y_i) | X_i] \right) \\
 &= \frac{1}{n} \sup_{f \in \mathcal{L}_1(B(0;1))} E \left[\sum_{i=1}^n (f(X_i) - f(Y_i)) | \mathbf{X} \right] \\
 &\leq \frac{1}{n} E \left[\left(\sup_{f \in \mathcal{L}_1(B(0;1))} \sum_{i=1}^n (f(X_i) - f(Y_i)) \right) | \mathbf{X} \right] \\
 &= \frac{1}{n} E [M_n^d | \mathbf{X}],
 \end{aligned} \tag{3}$$

hence from lemma 1, one has

$$E[W^d(m_n^d, \mu^d)] \leq \frac{1}{n} E[M_n^d] = O(n^{-\frac{1}{d}}).$$

□

Corollary 1. *Generally, let X_1, X_2, \dots, X_n be independent uniformly distributed random variables on the d -dimensional ball $B(0, \delta)$ with radius $\delta > 0$, and m_n^d denote the empirical measure*

$$m_n^d(y) = \frac{1}{n} \sum_{i=1}^n 1_{y=X_i},$$

and μ^d the uniform measure on $B(0, \delta)$. Then

$$E[W^d(m_n^d, \mu^d)] = \delta O(n^{-\frac{1}{d}}), d \geq 2.$$

Proof. Let $y = \delta t$ and denote

$$\bar{m}_n^d(t) = \delta^d m_n^d(\delta t) = \delta^d \frac{1}{n} \sum_{i=1}^n 1_{\delta t = X_i}, \bar{\mu}^d(t) = \delta^d \mu^d(\delta t),$$

then $\bar{m}_n^d, \bar{\mu}^d$ are empirical and uniform measures respectively about variable t on $B(0; 1)$. Especially one has

$$\begin{aligned}
 W^d(m_n^d, \mu^d) &= \inf_{\mu \in \Gamma(m_n^d, \mu^d)} \int_{B(0, \delta) \times B(0, \delta)} d(x, y) \mu(x, y) dx dy \\
 &= \inf_{\mu \in \Gamma(m_n^d, \mu^d)} \int_{B(0; 1) \times B(0; 1)} d(\delta t, \delta \tau) \mu(\delta t, \delta \tau) \delta^{2d} dt d\tau \\
 &= \inf_{\bar{\mu} \in \Gamma(\bar{m}_n^d, \bar{\mu}^d)} \int_{B(0; 1) \times B(0; 1)} \delta d(t, \tau) \bar{\mu}(t, \tau) dt d\tau \\
 &= \delta W^d(\bar{m}_n^d, \bar{\mu}^d).
 \end{aligned} \tag{4}$$

So

$$E[W^d(m_n^d, \mu^d)] = \delta O(n^{-\frac{1}{d}}), d \geq 2.$$

□

Theorem 2. Consider a Poisson process \mathcal{P} with intensity measure $(1 + \alpha_n)n \frac{dV_d}{|B(0; 1)|}$ on d -dimensional ball $B(0; 1)$, for some sequence $0 \leq \alpha_n \rightarrow 0$. Let $m_{\mathcal{P}}^d$ denote the empirical random measure with respect to \mathcal{P} , i.e.

$$m_{\mathcal{P}}^d(y) = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} 1_{y=p},$$

and μ^d the uniform measure on the above ball. Then

$$E[W(m_{\mathcal{P}}^d, \mu^d)] = O(n^{-\frac{1}{d}}).$$

Proof. Since the size of random variables about $m_{\mathcal{P}}^d$ is produced by mixed random process \mathcal{P} , so

$$E[W(m_{\mathcal{P}}^d, \mu^d)] = \sum_{k=1}^{\infty} E[W(m_{\mathcal{P}}^d, \mu^d) | |\mathcal{P}| = k] \mathbb{P}(Po((1 + \alpha_n)n) = k).$$

From Theorem 1, we know

$$E[W(m_{\mathcal{P}}^d, \mu^d) | |\mathcal{P}| = k] = O(k^{-\frac{1}{d}}).$$

Another hand, from lemma 1.2 in [4] one has

$$\begin{aligned}
 \mathbb{P}(Po((1 + \alpha_n)n) > (1 + \alpha_n)n + c\sqrt{(1 + \alpha_n)n \log n}) &\leq e^{c\sqrt{(1 + \alpha_n)n \log n} + ((1 + \alpha_n)n + c\sqrt{(1 + \alpha_n)n \log n}) \log \frac{(1 + \alpha_n)n}{(1 + \alpha_n)n + c\sqrt{(1 + \alpha_n)n \log n}}} \\
 &\leq e^{-\frac{c^2(1 + \alpha_n)n \log n}{2((1 + \alpha_n)n + c\sqrt{(1 + \alpha_n)n \log n})}} = O(n^{-\frac{c^2}{3}})
 \end{aligned} \tag{5}$$

where c is a constant determined later, and

$$\mathbb{P}(Po((1 + \alpha_n)n) < (1 + \alpha_n)n - c\sqrt{(1 + \alpha_n)n \log n}) \leq e^{-c\sqrt{(1 + \alpha_n)n \log n} + ((1 + \alpha_n)n - c\sqrt{(1 + \alpha_n)n \log n}) \log \frac{(1 + \alpha_n)n}{(1 + \alpha_n)n - c\sqrt{(1 + \alpha_n)n \log n}}} = O(n^{-\frac{c^2}{3}}). \tag{6}$$

Denote

$$a_n^{\pm} = [(1 + \alpha_n)n \pm c\sqrt{(1 + \alpha_n)n \log n}].$$

Then

$$\begin{aligned}
E[W(m_{\mathcal{P}}^d, \mu^d)] &= \sum_{k=1}^{a_n^- - 1} E[W(m_{\mathcal{P}}^d, \mu^d) \mid |\mathcal{P}| = k] \mathbb{P}(Po((1 + \alpha_n)n) = k) \\
&\quad + \sum_{k=a_n^-}^{a_n^+} E[W(m_{\mathcal{P}}^d, \mu^d) \mid |\mathcal{P}| = k] \mathbb{P}(Po((1 + \alpha_n)n) = k) \\
&\quad + \sum_{k=a_n^+ + 1}^{\infty} E[W(m_{\mathcal{P}}^d, \mu^d) \mid |\mathcal{P}| = k] \mathbb{P}(Po((1 + \alpha_n)n) = k) \\
&:= I_1 + I_2 + I_3.
\end{aligned}$$

Hence

$$I_1 = \sum_{k=1}^{a_n^- - 1} E[W(m_{\mathcal{P}}^d, \mu^d) \mid |\mathcal{P}| = k] \mathbb{P}(Po((1 + \alpha_n)n) = k) \leq 2\mathbb{P}(Po((1 + \alpha_n)n) < a_n^-) = O(n^{-\frac{c^2}{3}})$$

and

$$I_3 = \sum_{k=a_n^+ + 1}^{\infty} E[W(m_{\mathcal{P}}^d, \mu^d) \mid |\mathcal{P}| = k] \mathbb{P}(Po((1 + \alpha_n)n) = k) \leq 2\mathbb{P}(Po((1 + \alpha_n)n) > a_n^+) = O(n^{-\frac{c^2}{3}}).$$

The last part I_2 , one has

$$\begin{aligned}
I_2 &= \sum_{k=a_n^-}^{a_n^+} E[W(m_{\mathcal{P}}^d, \mu^d) \mid |\mathcal{P}| = k] \mathbb{P}(Po((1 + \alpha_n)n) = k) \\
&= \sum_{k=a_n^-}^{a_n^+} O(k^{-\frac{1}{d}}) \frac{e^{-(1+\alpha_n)n} ((1+\alpha_n)n)^k}{k!} \\
&\leq O\left(((1+\alpha_n)n - c\sqrt{(1+\alpha_n)n \log n})^{-\frac{1}{d}}\right) = O(n^{-\frac{1}{d}}).
\end{aligned}$$

Finally one has for some adjusted c

$$E[W(m_{\mathcal{P}}^d, \mu^d)] = O(n^{-\frac{1}{d}})$$

□

Corollary 2. Let $0 \leq \alpha_n \rightarrow 0$, $x \in \mathbb{R}^d$ and consider a Poisson process P with intensity measure $(1 + \alpha_n)n \frac{dV_d}{|B(0;1)|}$ on the d -dimensional ball $B(x; \delta)$ with the radius $\delta > 0$. Let m_x^d denote the empirical measure with respect to P , i.e.

$$m_x^d(y) = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} 1_{y=p},$$

and μ_x^d the uniform measure on $B(x; \delta)$. Then

$$E[W^d(m_x^d, \mu_x^d)] = O(n^{-\frac{1}{d}}).$$

Proof. As the proof in Theorem 2. First it has the size $|\mathcal{P}|$ a Poisson distribution on $B(x; \delta)$ with mean value $(1 + \alpha_n)n\delta^d$. So

$$E[W^d(m_x^d, \mu_x^d)] = E[\delta W^d(\bar{m}_x^d, \bar{\mu}_x^d)] = \delta E[W^d(\bar{m}_x^d, \bar{\mu}_x^d)] = \delta O((n\delta^d)^{-\frac{1}{d}}) = O(n^{-\frac{1}{d}}).$$

□

Conclusion Corollary 2 may be directly applied to get the Appendix A.3 in [1]. Next we hope to apply our method to analyze electrostatic approach problem.

A The lower bound of $E(M_n^d)$

Since $M_n^d = \inf_{\sigma} \sum_{i=1}^n \|X_i - Y_{\sigma(i)}\| \geq \inf_{\sigma} \sum_{i=1}^n \min_{j \leq n} \|X_i - Y_j\| = \sum_{i=1}^n \min_{j \leq n} \|X_i - Y_j\|$, one has

$$E(M_n^d | X) \geq \sum_{i=1}^n E(\min_{j \leq n} \|X_i - Y_j\| | X) \geq n \min_{x \in D} E(\min_{j \leq n} \|x - Y_j\|).$$

Let $B(x, t) = \{y \in R^d : \|x - y\| \leq t\}$, then $\frac{|B(x, t) \cap B(0; 1)|}{|B(0; 1)|} \leq \min\{t^d, 1\}$ and $P(\min_{j \leq n} \|x - Y_j\| \geq t) \geq (1 - t^d)^n$ for $t < 1$. Thus

$$E(\min_{j \leq n} \|x - Y_j\|) = \int_0^\infty P(\min_{j \leq n} \|x - Y_j\| \geq t) dt \geq \int_0^1 (1 - t^d)^n dt \stackrel{t=n^{-1/d}u}{=} n^{-1/d} \int_0^{n^{1/d}} (1 - u^d/n)^n du.$$

Finally by Fatou lemma, one has

$$\liminf_{n \rightarrow \infty} \frac{E(M_n^d)}{n^{1-\frac{1}{d}}} \geq \int_0^\infty e^{-u^d} du \geq 1 - \frac{1}{d+1}.$$

B The upper bound of $E(M_n^d)$

Set $r = \frac{1}{n^{\frac{1}{d}}}$ so $r \rightarrow 0$ as $n \rightarrow \infty$ and $|B(x, r)| = \frac{1}{n} \omega_d$, where $\omega_d = |B(0; 1)|$. Set

$$u(i, j) = \begin{cases} 1, & \text{if } \|X_i - Y_j\| \leq r \\ 0, & \text{otherwise} \end{cases}$$

Set

$$b(x) = \frac{|B(x, r) \cap B(0; 1)|}{|B(0; 1)|},$$

then $b(x) \leq \frac{1}{n} < 1$ if $n > 1$. First decomposed $\sum_{i \leq n} (f(X_i) - f(Y_i))$ as follows

$$\sum_{i \leq n} (f(X_i) - f(Y_i)) = \sum_{i \leq n} f(X_i) \sum_{j \leq n} u(i, j) - \sum_{j \leq n} f(Y_j) \sum_{i \leq n} u(i, j) + \sum_{i \leq n} f(X_i) (1 - \sum_{j \leq n} u(i, j)) - \sum_{j \leq n} f(Y_j) (1 - \sum_{i \leq n} u(i, j)),$$

so one has

$$\begin{aligned} \left| \sum_{i \leq n} (f(X_i) - f(Y_i)) \right| &\leq \left| \sum_{i \leq n} f(X_i) \sum_{j \leq n} u(i, j) - \sum_{j \leq n} f(Y_j) \sum_{i \leq n} u(i, j) \right| + \left| \sum_{i \leq n} f(X_i) (1 - \sum_{j \leq n} u(i, j)) - \sum_{j \leq n} f(Y_j) (1 - \sum_{i \leq n} u(i, j)) \right| \\ &:= I_1 + I_2. \end{aligned} \tag{7}$$

Further we will run two parts estimation of $E \sup_{f \in \mathcal{L}_1} I_1$ and $E \sup_{f \in \mathcal{L}_1} I_2$ and finally get

$$E(M_n^d) = E \sup_{f \in \mathcal{L}_1} \left| \sum_{i \leq n} (f(X_i) - f(Y_i)) \right| \leq E \sup_{f \in \mathcal{L}_1} I_1 + E \sup_{f \in \mathcal{L}_1} I_2 \leq n^{1-\frac{1}{d}} + 2dn^{1-\frac{1}{d}} + 4n^{1-\frac{1}{d}} + 8\sqrt{2}n^{\frac{1}{2}},$$

hence one has

$$\frac{E(M_n^d)}{n^{1-\frac{1}{d}}} \leq 1 + 2d + 4 + 8\sqrt{2}n^{\frac{1}{d}-\frac{1}{2}} = 5 + 2d + 8\sqrt{2}n^{\frac{1}{d}-\frac{1}{2}}.$$

B.1 The estimation of $E \sup_{f \in \mathcal{L}_1} I_1$

$$I_1 = \left| \sum_{i \leq n} f(X_i) \sum_{j \leq n} u(i, j) - \sum_{j \leq n} f(Y_j) \sum_{i \leq n} u(i, j) \right| = \left| \sum_{i \leq n} \sum_{j \leq n} u(i, j) (f(X_i) - f(Y_j)) \right| \leq r \sum_{i \leq n} \sum_{j \leq n} u(i, j).$$

Since f is Lipschitz,

$$E \sup_{f \in \mathcal{L}_1} I_1 \leq r \sum_{i \leq n} \sum_{j \leq n} Eu(i, j) = r \sum_{i \leq n} \sum_{j \leq n} E(E(u(i, j)|X_i)) = r \sum_{i \leq n} \sum_{j \leq n} E(b(X_i)) \leq r \sum_{i \leq n} \sum_{j \leq n} \frac{1}{n} = n^{1-\frac{1}{d}}.$$

Notice the estimation of I_1 may further optimize, refer to [5].

B.2 The estimation of $E \sup_{f \in \mathcal{L}_1} I_2$

Decomposed I_2 as follows

$$\begin{aligned} I_2 &\leq \left| \sum_{i \leq n} f(X_i) (1 - nb(X_i)) - \sum_{j \leq n} f(Y_j) (1 - nb(Y_j)) \right| + \left| \sum_{i \leq n} f(X_i) \left(nb(X_i) - \sum_{j \leq n} u(i, j) \right) \right| + \left| \sum_{j \leq n} f(Y_j) \left(nb(Y_j) - \sum_{i \leq n} u(i, j) \right) \right| \\ &:= I_{21} + I_{22} + I_{23}. \end{aligned}$$

We will run the estimation for these three parts and finally get

$$E \sup_{f \in \mathcal{L}_1} I_2 \leq E \sup_{f \in \mathcal{L}_1} I_{21} + E \sup_{f \in \mathcal{L}_1} I_{22} + E \sup_{f \in \mathcal{L}_1} I_{23} \leq 2dn^{1-\frac{1}{d}} + 4n^{1-\frac{1}{d}} + 8\sqrt{2}n^{\frac{1}{2}}.$$

B.2.1 The estimation of $E \sup_{f \in \mathcal{L}_1} I_{21}$.

Notice $\|f\|_{L^\infty[B(0;1)]} \leq 1$ and the value of $b(X_i)$ on $B(0; 1)$, one has

$$E \sup_{f \in \mathcal{L}_1} \left(\left| \sum_{i \leq n} f(X_i) (1 - nb(X_i)) \right| \right) \leq E \left(\sum_{i \leq n} |1 - nb(X_i)| \right) = \sum_{i \leq n} E(|1 - nb(X_i)|) \leq ndr = dn^{1-\frac{1}{d}},$$

so we can get

$$\begin{aligned}
E \sup_{f \in \mathcal{L}_1} I_{21} &= E \sup_{f \in \mathcal{L}_1} \left(\left| \sum_{i \leq n} f(X_i)(1 - nb(X_i)) - \sum_{j \leq n} f(Y_j)(1 - nb(Y_j)) \right| \right) \\
&\leq E \sup_{f \in \mathcal{L}_1} \left| \sum_{i \leq n} f(X_i)(1 - nb(X_i)) \right| + E \sup_{f \in \mathcal{L}_1} \left| \sum_{j \leq n} f(Y_j)(1 - nb(Y_j)) \right| \\
&\leq 2dn^{1-\frac{1}{d}},
\end{aligned}$$

B.2.2 The estimation of $E \sup_{f \in \mathcal{L}_1} I_{22}, E \sup_{f \in \mathcal{L}_1} I_{23}$.

This part estimation is difficult and one may use the convolution decomposition to impose f on small areas, then there will has the below estimation

$$E \sup_{f \in \mathcal{L}_1} I_{22} = E \sup_{f \in L_1} I_{23} = E \sup_{f \in \mathcal{L}_1} \left(\left| \sum_{i \leq n} f(X_i)(nb(X_i) - \sum_{j \leq n} u(i, j)) \right| \right) \leq 2n^{1-\frac{1}{d}} + 4\sqrt{2}n^{\frac{1}{2}}.$$

First we assume f an indicator function on some set A , and estimate $E \left(\left| \sum_{i \leq n} 1_A(X_i)(nb(X_i) - \sum_{j \leq n} u(i, j)) \right|^2 \right)$.

So one has

$$E \left(\left| \sum_{i \leq n} 1_A(X_i)(nb(X_i) - \sum_{j \leq n} u(i, j)) \right|^2 \right) = E \left(\sum_{i, i' \leq n} \sum_{j, j' \leq n} 1_A(X_i)(b(X_i) - u(i, j)) 1_A(X_{i'})(b(X_{i'}) - u(i', j')) \right).$$

Discussing as different cases of i, i', j, j' , there has

$$E \left(\left| \sum_{i \leq n} \sum_{j \leq n} 1_A(X_i)(b(X_i) - u(i, j)) \right|^2 \right) \leq n^2(n-1) \frac{|A|}{|B(0; 1)|} \frac{1}{n^2} + n^2 \frac{|A|}{|B(0; 1)|} \frac{1}{n} \leq 2n \frac{|A|}{|B(0; 1)|}. \quad (8)$$

Second we consider convolution. Let g is bounded Lipschitz in R^d , h a bounded support function in R^d , and $g * h(x) = \int_{R^n} g(t)h(x-t)dt$ to estimate the expectation of

$$\left| \sum_{i \leq n} g * h(X_i) \sum_{j \leq n} (b(X_i) - u(i, j)) \right| = \left| \int_{R^n} g(t) \sum_{i \leq n} \sum_{j \leq n} h(X_i - t)(b(X_i) - u(i, j)) dt \right|.$$

That is

$$\begin{aligned}
&E \left(\left| \sum_{i \leq n} g * h(X_i) \sum_{j \leq n} (b(X_i) - u(i, j)) \right| \right) \\
&= E \left(\left| \int_{R^n} g(t) \sum_{i \leq n} \sum_{j \leq n} h(X_i - t)(b(X_i) - u(i, j)) dt \right| \right) \\
&\leq \|g\|_{L^\infty} E \left(\int_{R^n} \left| \sum_{i \leq n} \sum_{j \leq n} h(X_i - t)(b(X_i) - u(i, j)) \right| dt \right) \\
&\leq \|g\|_{L^\infty} \left(\int_{B(0; 1)^{2n}} \int_{R^n} \left| \sum_{i \leq n} \sum_{j \leq n} h(X_i - t)(b(X_i) - u(i, j)) \right|^2 \rho dt d\sigma \right)^{\frac{1}{2}} \times \left(\int_{B(0; 1)^{2n}} \int_{R^n} 1_{x-t \in \text{supp } h}(x, t) \rho dt d\sigma \right)^{\frac{1}{2}} \\
&= (\text{supp}(h))^{\frac{1}{2}} \|g\|_{L^\infty} \left(\int_{R^n} E \left(\left| \sum_{i \leq n} \sum_{j \leq n} h(X_i - t)(b(X_i) - u(i, j)) \right|^2 \right) dt \right)^{\frac{1}{2}}.
\end{aligned}$$

Further set $h(x) = c_0 1_A(x)$, there is by formula (8)

$$\int_{R^n} E\left(\left|\sum_{i \leq n} \sum_{j \leq n} h(X_i - t)(b(X_i) - u(i, j))\right|^2\right) dt = \int_{R^n} \int_{B(0; 1)^{2n}} \left|\sum_{i \leq n} \sum_{j \leq n} c_0 1_A(X_i - t)(b(X_i) - u(i, j))\right|^2 \rho d\sigma dt \leq 2nc_0^2|A|,$$

hence one has the below estimation of g convolution with characteristic function

$$E\left(\left|\sum_{i \leq n} g * (c_0 1_A)(X_i) \sum_{j \leq n} (b(X_i) - u(i, j))\right|\right) \leq |A|^{\frac{1}{2}} c_0 (2|A|n)^{\frac{1}{2}} \|g\|_{L^\infty} = \sqrt{2} c_0 |A| n^{\frac{1}{2}} \|g\|_{L^\infty} \quad (9)$$

Finally We decompose f into a sum of some well-defined convolutions to estimate these parts. Since a Lipschitz function f , in $B(0; 1) \subset R^d$ with $f(0) = 0$, may extend to a whole Lipschitz function on R^d with $\|f\|_{L^\infty} \leq 1$. So we may consider function $f \in \mathcal{L}$ and decompose $f = \sum_{l=1}^{q+1} f_l$, where

$$h_l = \begin{cases} |B(0; 1)|^{-1} (2^l r)^{-d} & x \in B(0, 2^l r) \\ 0 & \text{otherwise} \end{cases},$$

$f_1 = f - f * h_1$, $f_2 = f * h_1 - f * h_2 * h_1$, \dots $f_q = f * h_{l-1} * \dots * h_1 - f * h_l * h_{l-1} * \dots * h_1$, $f_{q+1} = f * h_q * \dots * h_1$, and q denoted by $2^{q+1} \frac{\eta^d}{n} = 2^{q+1} r \geq 1 > 2^q r = 2^q \frac{\eta^d}{n}$. So

$$E \sup_{f \in \mathcal{L}_1} I_{22} = E\left(\sup_{f \in \mathcal{L}} \left|\sum_{i \leq n} \sum_{l=1}^{q+1} f_l(X_i) \sum_{j \leq n} (b(X_i) - u(i, j))\right|\right) \leq \sum_{l=1}^{q+1} E\left(\sup_{f \in \mathcal{L}} \left|\sum_{i \leq n} f_l(X_i) \sum_{j \leq n} (b(X_i) - u(i, j))\right|\right).$$

For first part, one has

$$E\left(\left|\sum_{i \leq n} f_1(X_i) \sum_{j \leq n} (b(X_i) - u(i, j))\right|\right) \leq \sum_{i \leq n} \|f - f * h_1\|_{L^\infty} E\left(\left|\sum_{j \leq n} (b(X_i) - u(i, j))\right|\right) \leq n2r,$$

where we omit the computing process of $E\left(\left|\sum_{j \leq n} (b(X_i) - u(i, j))\right|^2\right) \leq 1$, further $E\left(\left|\sum_{j \leq n} (b(X_i) - u(i, j))\right|\right) \leq 1$.

For the expectation about $f_l = (f - f * h_l) * h_1 * \dots * h_{l-1}$ with $2 \leq l \leq q$, one has from formula (9)

$$E\left(\left|\sum_{i \leq n} f_l(X_i) \sum_{j \leq n} (b(X_i) - u(i, j))\right|\right) \leq \sqrt{2} |B(0; 1)|^{-1} (2^{l-1} r)^{-d} |\text{supp}(h_{l-1})| n^{\frac{1}{2}} \|(f - f * h_l) * h_1 * \dots * h_{l-2}\|_{L^\infty} \leq \sqrt{2} n^{\frac{1}{2}} 2^l r$$

For the last part expectation about $f_{q+1} = f * h_1 * \dots * h_{q-1} * h_q$, one has by (9) again

$$E\left(\left|\sum_{i \leq n} f_{q+1}(X_i) \sum_{j \leq n} (b(X_i) - u(i, j))\right|\right) \leq \sqrt{2} |B(0; 1)|^{-1} (2^q r)^{-d} |\text{supp}(h_q)| n^{\frac{1}{2}} \|f * h_1 * \dots * h_{q-1}\|_{L^\infty} \leq \sqrt{2} n^{\frac{1}{2}} 2^{q+1} r.$$

Summing all part estimations, one has

$$E \sup_{f \in \mathcal{L}_1} I_{22} \leq n2r + (\sqrt{2} n^{\frac{1}{2}} 2^2 r + \dots + \sqrt{2} n^{\frac{1}{2}} 2^{q+1} r) \leq 2n^{1-\frac{1}{d}} + 4\sqrt{2} n^{\frac{1}{2}}$$

References

- [1] Hoorn P. van der , Lippner G., Trugenberger C., Krioukov D. *Ollivier-Ricci curvature convergence in random geometric graphs.* Physical Review Research 3, 013211 (2021).
- [2] Kuhn H. W. *The Hungarian method for the assignment problem.* Naval Res. Logistics Quarterly, Volume2, Issue1-2, March 1955, Pages 83-97.
- [3] Papadimitriou C. H., Steiglitz K. *Combinatorial Optimization, Algorithms and Complexity.* Prentice-Hall, Englewood Cliffs, N.J. (1982).
- [4] Penrose M. *Random geometric graphs.* Oxford university press, 2003.
- [5] Talagrand M. *Matching random samples in many dimensions.* The Annals of Applied Probability 1992, Vol. 2, No. 4, 846-856.
- [6] Veronika Kralj-Iglič, Aleš Iglič, *A Simple Statistical Mechanical Approach to the free Energy of the Electric Double Layer Including the Excluded Volume Effect.* Journal de Physique II, Volume 6, Issue 4, April 1996, pp.477-491.
- [7] Villani C., *Topics in optimal transportation.* Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003.